4-2-2004

Georeferencing Digital Camera Images Using Internal Camera Model

Alok Nagdev
University of South Florida

Georeferencing Digital Camera Images Using Internal Camera Model

by

Alok Nagdev

A thesis submitted in partial fulfillment
of the requirements for the degree of
Master of Science in Computer Science
Department of Computer Science and Engineering
College of Engineering
University of South Florida

Major Professor: Murali Varanasi, Ph.D.
John Brock, Ph.D.
Dmitry Goldgof, Ph.D.

Date of Approval:
April 2, 2004

Keywords: image correlation, attitude, image projection, camera calibration, least squared estimation

ACKNOWLEDGEMENTS

I would like to thank Dr. Varanasi, my major professor, for many ideas and suggestions, for great encouragement, and for giving me the opportunity to work on this project with the US Geological Survey. I would sincerely like to thank Dr. John Brock, my supervisor at USGS, for his insights and wealth of knowledge, which has contributed immensely towards my thesis. Dr. Brock has patiently guided me and I have gained tremendous wisdom from the discussions we have had, which go beyond the realms of a thesis. I also thank him for painstakingly reviewing and critiquing my thesis writeup. I would also like to thank Dr. Goldgof for all comments and help, for sharing his expert competence and for great support throughout. My sincere thanks to US Geological Survey for their financial support. A special thanks to my family and friends for supporting me and being there for me when I needed them most.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

GEOREFERENCING DIGITAL CAMERA IMAGES USING INTERNAL CAMERA MODEL

Alok Nagdev

ABSTRACT

The NASA Airborne Topographic Mapper (ATM) is a laser scanning instrument which is used mainly to collect dense topographic data over much of the conterminous US coastline. The inclusion of two digital cameras in consonance with the ATM instrument now gives 3-band (RGB) imagery apart from a very rich topographic data. This imagery, in its crude form, has limited applications due to it being not georeferenced and having a heavy camera lens distotions. As thesis, a processing system - Park-View - is developed to bring this imagery into a more suitable format for the scientists for alalytical and interpretational purposes. Park-View utilizes the well gridded elevation data from layer four of another processing system called LaserMap for georeferencing the digital camera images. Camera lens behavior is modeled using a 2D grid image and all of its intrinsic parameters ascertained. These parameters are then incorporated into correcting the lens distortions of georeferenced images. Errors in time-stamping of images and in the mounting angles of the camera are calculated using well known tie-points. Georeferenced images can be stored either in GeoTiff format or jpeg format. Individual images can be georeferenced or put in a mosaic form with the mosaic color equalized for adjoining images. Park-View also provides the main GUI displaying the entire surveyed area, mapper GUI for a batch processing of all the images and a display window for displaying georeferenced images or mosaics. Additional capabilities could be added to the processing system for performing some specific image processing operations on the images such as edge detection and image enhancement.

CHAPTER 1

INTRODUCTION

In recent years, technological advances have changed the way geographic analyses are done. Increasingly, computers are used to automate aspects of cartography and remote sensing, producing data that are easily integrated into a Geographic Information System (GIS).

Use of computational techniques have considerably reduced the time taken to process remotely sensed data. Various software processing techniques have evolved to process remotely sensed data to bring it to its interpretational level. The remotely sensed data, over the years, has come a long way from the days of analog to digital form of data. After the advent of computers and associated increase in the processing speed of the computers, digital remote sensing has taken over the conventional remote sensing technique in terms of processing speed. Various remote sensing techniques are combined with each other to enhance the capability of each other giving a new dimension to the field of remote sensing. The information collected through all these methods about the earth surface is growing in size as newer techniques are evolving, hence necessitating the need to efficiently handle and process this data. Newer methods need to be applied to process such huge amounts of data. This thesis is an effort in that direction.

The ongoing research in the field of geological changes as carried on by the United States Geological Survey (USGS) uses many diverse remote sensing techniques, which consequently are supported by different software processing systems. Apart from this USGS also has vast earth and biological data holdings. One related series of research is going on with mapping the US coastline using the LIDAR data. This topographic data then is used in applications such as creating the Digital Elevation Models (DEM), coastline changes, study and classification of vegetation, and studies on the coral reefs [4] [5]. Recently the

aircrafts carrying the LIDAR instruments were also equipped with the digital cameras for taking photographs of the surveyed areas along with the LIDAR data. For digital aerial images to be useful to the scientists, these images have to be georeferenced with respect to some existing reference system. Other agencies like National Aeronautics and Space Administration (NASA) and National Oceanic and Atmospheric Administration (NOAA) are instrumental in providing relevant information ranging from the base level data to a high end processed data and images.

The theme of this thesis is combining of the LIDAR data with the image data involving some computer vision techniques to georeference the digital camera images and the subsequent development of the related processing system. This thesis contrasts two different methods to carry out the task of georeferencing the digital camera images, and compares the results. The first approach uses the simple image processing techniques whereas the second approach uses computer vision methods.

1.1 Mapping Techniques

There are several different mapping techniques that biological and coastal scientists use. These techniques range from the age old T-sheet mapping to the more recent LIDAR remote sensing and multispectral imaging. The fact is that sometimes the analysis is carried without georeferencing the data (for e.g. aerial photographs) or most of the times with georeferencing. Quantification of the geological changes can be more accurately determined if the data sets in comparison are precisely georeferenced with some reference frame.

1.1.1 T-Sheets

The T-sheets (Topographic Sheets) are issued by the National Oceanic and Atmospheric Administration's (NOAA) National Ocean Service (NOS), formerly known as U.S. Coast and Geodetic Survey (USC&GS). T-sheets are detailed survey maps that were produced to provide coastlines for use on navigation charts. The information content of T-sheets varies. At a minimum each T-sheet contains following: the mean high water line as derived from field survey; survey marker locations used to provide control for the field work; a coordinate

system formed by longitude and latitude lines; and graphic representation of the vegetation inland from the coast.

As for research involving quantitative analysis of changes in configuration of the shorelines and its change in position over time, it is essential that the historical T-sheets are converted to a digitized form for the comparison purposes. So, in general the use of T-sheets for comparison studies introduces several potential errors, inherent to it as well as due to conversion process, to the coastal mapping process. The severity of these errors depends on the accuracy standards met by each map and on the physical changes in each map since publication [25].

1.1.2 Aerial Photographs

Aerial photography is extremely useful in context of mapping. It has been extensively used in the preparation of maps in which cartographers and planners take a detailed measurement from the aerial photographs. Besides, trained users and interpreters utilize them for finding land-use, land cover classification, environmental conditions among other things.

Aerial photographs are not really maps, even though they present a "birds-eye" view of the earth's surface. Maps are orthogonal representations of the earth's surface with a certain geometrical and directional accuracy (generally imposed by National Map Accuracy Standards, according to the scale of the map).

History of photography dates back to 1827 when Joseph Nicephore Niepce took the first picture of nature where the exposure time was 8 hours and the emulsion was bitumen of Jedea. Later in 1858 the first known aerial photograph is taken from a captive balloon from an altitude of 1,200 feet over Paris by Gaspar Felix Tournachon Nadar. Since then many advances have taken place in the field of remote sensing and aerial photography. There are now not just three band photography that is used but also multi-spectral photography emerging that supports up to 48 bands. First photograph of the curvature of the earth was taken from a free balloon at an altitude of 72,000 ft. by Captain Albert W. Stevens in 1936.

Even though the aerial photographs are a very common source of data for the interpretational purposes they cannot be used for quantifying the coastal changes for example. Aerial photographs display a high degree of radial distortion due to the camera lens distortion and the camera film deformation. Apart from this camera pitch, roll, and yaw create deformations in object space. As the digital cameras do not have emulsive films they do not pose the distortion due to deformation of the film.

1.1.3 Digital Orthophoto Quadrangles (DOQ)

A Digital Orthophoto Quadrangle (DOQ) is a rectified digital image of an aerial photograph with distortions and displacements removed and corrected for aircraft pitch, roll, yaw and altitude, camera tilt and terrain relief, thus complying to the properties of an orthographic projection. DOQs are usually generated by scanning the aerial photograph converting it into a digital format and applying a rectifying process to rectify it. A DOQ, is therefore, a computer generated image of an aerial photograph that transforms the image characteristic of a photograph into the geometric qualities of a map.

The aerial photographs are digitized and processed to produce 3.75 x 3.75 minute DOQs with 1 meter ground resolution and mapped to 1:12,000 scale accuracy specifications. Projections are based on the North American Datum of 1983 (NAD 83) and the images are produced on the Universal Transverse Mercator (UTM) projection. Normally each DOQ image is over-edged by 50 to 300 meters to facilitate the tonal matching for mosaicing of the images.

A DOQ may be used as a source or as a background layer in a Geographic Information System (GIS) to collect, revise, and review other digital data. DOQs can also be combined with Digital Elevation Model (DEM) images or some other type of image to produce a more informative image for the cartographic or analysis or investigation purposes.

1.1.4 Satellite Imagery

Remote sensing involves gathering data and information about physical world by detecting and measuring radiations, particles, fields emanating from objects located beyond

the immediate vicinity of the sensor devices. Remotely sensed images are usually categorized according to the altitude of the aircraft or spacecraft and the characteristic of the sensor used to generate images. Satellite remote sensing thus is the imagery captured from a satellite stationed sensor. Satellite images can capture information regarding earth surface, vegetation, land-water delineation as well as about the galaxy and beyond. Satellites, including manned space-crafts collect imagery from hundreds of miles above the earth's surface. Because of the altitude of the satellite sensors the image resolution is low about 1km by 1km as compared to the low-altitude photographs that have as high resolution as 10cm by 10cm on ground. In contrast satellite imagery covers large regions, suitable for the meteorological applications that require coarse resolution. Still, it is the space systems including satellites that have contributed enormously towards regional and global geophysical surveys. Firstly it is very difficult and costly to organize and conduct ground and aerial surveys at that large scale and secondly coordinating those individual surveys would be a real tough task.

United States launched its first experimental ERTS-1 (Earth Resources Technology Satellite, which was later renamed as LANDSAT) for global coverage of Earth's land masses. A program that was initiated by the U. S. Department of Interior and NASA, it was the first unmanned satellite designed to acquire the earth resource data on a repetitive, multispectral basis. Later LANDSAT 2 through 5 were launched during a period from 1975 to 1984. These satellites were equipped with sensors to collect data from a range of wavelengths of electromagnetic energy emitted or reflected from Earth's surface. In principle, a remote sensing unit can sense any range of wavelengths of electromagnetic spectrum. Typically wavelength ranges between 0.4μm to 12μm (visible/infrared range) and between 30mm to 300mm (microwave range) are used to collect data. Visible satellite images record visible light from the sun reflected back to the satellite from land surfaces, oceans or cloud surfaces. Infrared satellite images record invisible infrared radiation emitting directly from the land surfaces, oceans or cloud surfaces. The warmer the object is more is the infrared energy emitted by it, so this image characterizes the image features' temperatures.

LANDSATs 1 through 5 carried a versions of sensors called Multispectral Scanner (MSS), which simultaneously collected data from four broad bands of the electromagnetic spectrum, ranging from visible green through near-infrared wavelengths. The images provided by the MSS sensor had a resolution of about 80 meters on Earth's surface. In addition to MSS, LANDSAT 4 and 5 also carried the Thematic Mapper (TM) sensor. In addition to the higher resolution on the ground of 30 meters, TM sensor also records more number of bands than the MSS, providing more detailed spectral information. Latest in a series of earth monitoring satellites is Landsat 7, launched in April 1999. The sensor used on Landsat 7, Enhanced Thematic Mapper (ETM+), replicates its successor TM sensor but also provides new features such as a panchromatic band with 15 meter spatial resolution and a thermal infrared channel with 60 meter spatial resolution. These new features make this instrument more versatile and efficient for global change studies, land cover monitoring and assessment, and large area mapping than its predecessors.

The Advanced Very High Resolution Radiometer (AVHRR) sensor is aboard one of meteorological satellites operated by National Oceanic and Atmospheric Administration (NOAA). The AVHRR data are collected in 5 spectral bands in the visible, near-infrared, and thermal-infrared portions of the electromagnetic spectrum at a resolution of 1.1 Km. The data from AVHRR is being used for applications such as drought monitoring, fire fuel assessment, and land surface characterization.

There are other commercial imaging agencies that provide geographic information from optical and Earth observation satellites, compatible with all image processing and GIS systems. The range of products that these agencies provide include geocoded and ortho-corrected images, DEMs and land classification datasets. SPOT image by SPOT Image Corporation and IKONOS by Space Imaging Company are examples of such satellite imageries. IKONOS images have a swath width of 11Km at nadir and have a ground resolution of 1m for panchromatic and 4m for multispectral data. There are other satellite imaging systems that use active Radio Detection and Ranging (RADAR) sensor operating in the microwave region. Radar can see objects through such impervious conditions as darkness, haze, fog, rain and snow for its wavelengths are much longer than those of visible or in-

frared waves. Synthetic Aperture Radar is a common and very popular technique in radar imaging that achieves a very high resolution.

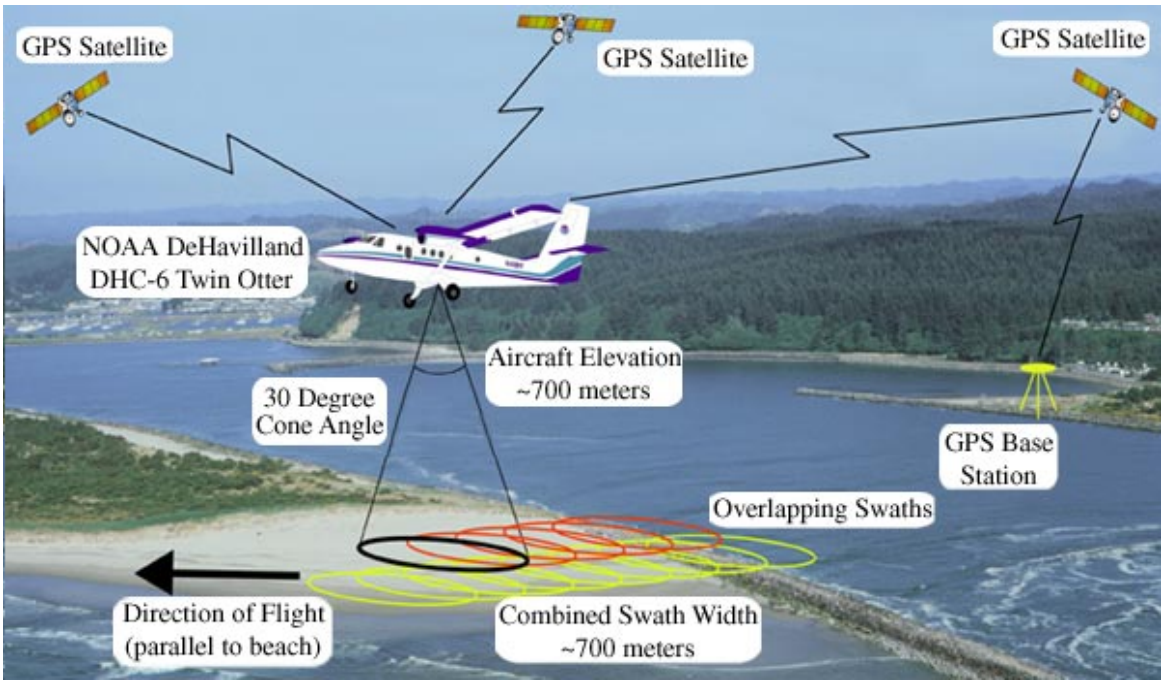


Figure 1.1 Diagram Showing the Elliptical Scan Pattern of NASA's ATM Operated from a NOAA Twin Otter. Digital Photo Camera is not Shown. (source: USGS)

1.1.5 Airborne Remote Sensing Systems

As suggested by its name, in airborne remote sensing, remote sensing instruments are mounted on an aircraft facing nadir (see Figure 1.1). Airborne remote sensing is good for collecting data on a small region basis because it provides very high resolution on the Earth's surface as compared to the satellite imaging. It also gives the choice of varying data density over a region of interest, but has the disadvantage of high cost per unit area of survey, the low coverage area as compared to the satellite imaging systems and the discontinuous nature of collecting data. Despite all this airborne remote sensing systems are frequently used for localized high resolution applications such as shoreline changes, coastal geology and coral reef studies. The different types of sensors that an aircraft carry include, but are not limited to, radar sensor and LIDAR (Light Detection and Ranging)

sensor and even cameras for aerial photography. Radar and LIDAR systems are called active systems as they send the microwave radiation (laser beam in case of LIDAR sensor) towards Earth's surface and measure the time taken for the radiation to return back to the receiver. This time is then used to calculate the distance of the reflected surface from the aircraft. Distance information combined with the GPS (Global Positioning System) locates a point in a known coordinate system. Also the albedo of an object could be ascertained if the amount of energy transmitted and the amount of energy reflected back is known, albeit atmospheric scattering, refraction and diffraction may cause a little bit of inaccuracy. Lower the altitude of the aircraft lower is the effect due to these factors.

1.2 Overview of Park-View

With the huge amount of data being collected during each survey involving NASA's ATM (Airborne Topographic Mapper) instrument, it became necessary to have this data processed quickly for analysis purposes for the scientists. One such processing system, LASERMAP [26], already exists for processing the topographic LIDAR data. With the inclusion of a digital photographic camera within the ATM instrument, need arose for processing these images as well. This means that these images are to be georeferenced and scaled properly before any constructive analysis could be made using them. In a broader sense georeferencing is nothing but aligning digital aerial camera images (individually or as a mosaic) with respect to some coordinated system with a proper scale. For orthogeoreferencing the image projection is assumed to be a parallel projection optics.

Park-view is a processing system that automatically orthogeoreferences the digital aerial camera imagery and also provides the capability to view the corrected images and mosaics. Different components of Park-View are

- Viewer Graphical User Interface (GUI): Gives a graphic representation of survey area and is used for preparing image mosaics.
- Mapper GUI: Used for preparing individual georeferenced images in a batch mode.
- Map Display Window: Used for displaying georeferenced images and maps.

Two completely different type of instruments (both LIDAR) gather information that can be used with the proposed approach. Airborne Topographic Mapper (ATM) instrument calculates the elevation data between the specific latitude and longitude ranges. This ATM device is coordinated with ground-based GPS receivers and an Inertial Navigation Unit (INU). Using differential kinematic GPS techniques [21], the position of the aircraft can be determined at an accuracy of within 5 centimeters [3]. Data collected using the ATM instrument is then run through various levels of processing before a DEM image grid is output. EAARL instrument works in a similar fashion with the difference from ATM being that it also records multiple returns from the object surface. This is very useful in terms of detecting water turbidity constants, land-cover classification, etc. As these instruments are collecting data in a survey, mounted cameras within the aircraft capture aerial images. The information regarding the time-stamping of the images is passed on to a control unit. Where control unit gathers information from all resources. ATM instrument carries a Kodak DC4800 camera with it having a resolution of 2160 by 1440 pixels, whereas AXIS2100 camera attached to EAARL instrument has a screen resolution of 352 by 240 pixels. Both of these cameras are 3-band (red, green and blue channel) cameras.

For preparing a orthogeoreferenced image/mosaic, flight track and aircraft attitude data is gathered for the specified region. This flight track data is then associated with each of the image taken of that region. The DEM data for the region is extracted and image back-projection is applied to each pixel point in the region to find the corresponding pixel intensity values. Before image back-projection is applied camera is calibrated for the camera's intrinsic parameters.

Time-stamping of the images taken is not perfect due to the unknown amount of time taken by camera to transfer the image back to the system where system records the time the image was taken as the current system time. Furthermore the mounting of the camera might have mounting biases from the pitch, roll and yaw angles calculated for the aircraft by the Inertial Navigation Unit (INU). As the aircraft moves at a speed of more than 40 meters per second, even a fraction of error in the time-stamping causes huge errors on the ground. Pitch, roll and yaw errors on ground are proportional to the altitude of

the aircraft. These errors too translate to huge errors on the ground for the slightest error in these angles. To correct these errors a few ground control points, whose ground position is well known, are chosen in an image and these points are then back-projected on the ground. The over-determined system of nonlinear equations using these GCPs is constructed and is solved for six unknowns using least square approximation. These estimated parameters are used to calculate time-stamping and mounting bias errors. These biases apply to all the images taken during a single mission, hence all the images of that mission could be orthogeoreferenced using the same camera model, without need for having ground control points for each and every image in the data set. This method is opposed to the classical orthogeorectification methods where every image to be orthogereferenced has to have ground control points within it. Although this flexibility is not for free and is at the cost of accuracy lost in the process.

1.3 Thesis Organization

Remaining chapters of the thesis are organized as follows: Chapter 2 will discuss some of the related work done in the field of orthogeorectification and image back-projection. Chapter 3 will explain the proposed method for the orthogeorectification of digital camera images. Chapter 4 will present and analyze the results. Chapter 5 will conclude and present the future direction of our research work.

CHAPTER 2

RELATED WORK

We will discuss some of the existing work in the area of georectification of camera images as well as some related work in the field of processing LIDAR data, which is a required input in our approach. The major emphasis would be on the photogrammetric techniques used in georeferencing the image data. Typically in georeferencing we apply three or more control points with known map coordinates to the object that is being georeferenced in order to tie the geographic location of cells in raster objects and elements in vector to a coordinate system and map projection of our choice. However, we can georeference objects without using multiple control points in those instances when we are unable to obtain enough information to locate and apply multiple control points. We can create a simple georeference for a raster object or an implied georeference for vector. These georeferences have limitations in their accuracy but are adequate for most uses that we might have.

A simple georeference can be created quickly for a raster object using only one control point with known coordinates, the cell size, and the projection of the object. A simple georeference is not as accurate as the results of other georeference methods but is adequate if we are concerned primarily with displaying the object. An implied georeference for a group of objects can be created when we do not have available the information needed for any of the other georeferencing methods. An implied georeference does not use control points but instead is based on a coordinate system and map projection that we choose for the group. An implied georeference allows a group of objects to be displayed correctly together but not with other objects.

2.1 Matching Aerial Image with Existing Orthophoto

The orientation of the images is the first step in the production of orthoimages. If orthoimages have to be renewed frequently, the existing orthoimages can be used for a fully automatic orientation process. The determination of the outer orientation of an aerial image requires camera data, approximate data for the outer orientation, and control points with known ground coordinates and observed image coordinates. For an automatic determination, the control points are replaced by control structures and a correspondence between orthoimage structures and aerial image structures has to be found at various positions. The orthoimage is georeferenced which means that each pixel has XY-coordinates in a reference system. From an existing digital elevation model (DEM) Z-values can be interpolated so that each pixel of the orthoimage is known by XYZ-coordinates.

Orientation parameters are calculated using the intensity parameter derived by converting the RGB image into HSI format. The relation between image and the orthoimage is given by well known collinearity equations.

Observation data can be established if parts of the orthoimage are matched with the aerial image. The image parts have to be sufficiently large, well distributed over the image area and with a good structure in order to find a good correlation. A small image part called a template is moved pixel by pixel over a somewhat larger image part (search area), and at each location a correlation coefficient is calculated after:

$$r = \frac{(I_1 - {I_1}^1) \cdot (I_2 - {I_2}^1)}{\sqrt{(I_1 - {I_1}^1)^2 \cdot (I_2 - {I_2}^1)^2}} \tag{2.1}$$

Where,

r is correlation coefficient

I_1 is the intensity value of a pixel in the template area

I_2 is the intensity value of a pixel in the search area

${I_1}^1, {I_2}^1$ average value of the template or search area respectively

The position where the correlation coefficient attains maximum value is the position of best fit. The result of the matching is ground coordinates for a pixel of each search area. With a sufficient number (about 10-25) of these observables the outer orientation can now be calculated using robust adjustment procedures. This leads to corrections of the outer orientation. The whole process is repeated leading to new corrections and further improved orientation data. The resolution of the search and template area is changed in any new orientation step [17].

2.2 Aerial Triangulation

Aerial triangulation is used to create additional ground control points at key locations on the aerial photograph for the setup of a single, pair, or multiple photographs for digital mapping, digital orthophoto, cross sections, profiles and volume inventory.

Before the aerial survey is flown, targets may be set on existing published ground control monuments or new locations may be selected and targeted. Even after the aerial survey has been flown some photo identifiable points may be selected and used. The density of the ground control points is determined according to the mapping requirements.

The triangulation process begins by laying out the photographs for each flight line and selecting locations for the minor control points on each photograph. Minor control points are selected down the center of each photograph and are given a unique number in accordance with their position on the photograph and location in the block.

Coordinate data files are prepared either by analog, analytical or softcopy methods for measuring minor control points along with any surveyed control points that fall on each photograph. This created file along with the ground control file becomes the input for the block adjustment program.

The block adjustment creates ground coordinates and elevations for all generated minor control points. These points are then used as control points to setup the photographs with the identical tips and tilts that were present in the aircraft at the time the aerial photography was flown.

2.3 Indirect Georeferencing

The general problem of georeferencing is that at the time of exposure no knowledge on the image position and orientation is available and hence the image vector is pointing to an arbitrary point in the object space. To map the image data to a suitable geographic coordinate frame, the position (X_0, Y_0, Z_0) and orientation (θ, ϕ, ω) (also known as the six external orientation parameters) of the sensor have to be known. If the image vector is corrected using these parameters then the data points on the image can be transformed into the global object reference frame points.

Since quite a long time the indirect image orientation remains the most favored approach for determining the orientation of traditional analogue (i.e. photogrammetric) cameras or the image based frame sensors. A bundle adjustment process is used to determine the exterior orientation of each image which are assumed to be unknowns. If no additional sensors are available for the positioning and orientation of the image sensor this remains the only way to determine the sensor position and the sensor orientation. The six unknown orientation parameters are then estimated using a set of ground control points and their image coordinates in a process called bundle adjustment. For the orientation of multiple images a process called aerial triangulation [32] is used wherein adjacent images are connected by a set of points having the same relative position, also called the tie points. In addition to the orientation parameters, object point coordinates and if desired additional calibration parameters could be estimated. If the mathematical model for the physical process of image formation is assumed as the standard central perspective then the georeferencing is based on the collinearity equations. This model associates the image coordinates to the object coordinate system. The interior orientation parameters define the camera geometry (i.e. principal point, focal length of the imaging lens etc.). Time dependent exterior orientation parameters reflect the camera position with respect to the global object coordinates.

Camera internal orientation (also called the camera's intrinsic parameters) measurement is usually carried out via laboratory calibration before it could be used in the bundle adjustment procedure. Even if the camera internal calibration is carried out in the labo-

ratories more advanced mathematical models could be used to improve the object point accuracy. This can be done by introducing additional parameters, apart from the six external orientation parameters, into the mathematical model. Introducing more parameters necessitates more Ground Control Points (GCP) for the eventual estimation of the parameters but in case of Aerial Triangulation (GCP) this gives better object point accuracy. Different mathematical models exist for the internal orientation estimation that correct for the systematic errors [6], [12]. Because of a strong correlation between different parameters the estimated external orientation parameters may not reflect the true physical exterior orientation parameters and are optimal values in the sense of object reconstruction only [30], [10]. This is because it compensates for all the systematic errors that are left behind by the self-calibration. This behavior is tolerable when AT process is used, as there's no need for the knowledge of the true physical orientation parameters but when direct georeferencing is used this might be a different problem altogether.

For satellite-borne sensors, due to their smooth trajectory, exterior orientation can be modeled as a polynomial function depending on time, with the physical properties of satellite's orbit as constraints (see e.g. [23]). A sufficient number of well distributed GCPs are required to solve the system. For airborne sensors, where the trajectory is unpredictable, modeling of the trajectory becomes harder and hence the need for direct georeferencing becomes essential.

Because of the different CCD array alignments used in different sensors mathematical models used for indirect georeferencing also vary. There is a wide class of linear CCD array sensors that acquire images in a pushbroom mode (e.g. ADS40, DPA and WAAC from DLR, AirMISR from NASA all airborne and SPOT from CNES, IRS from ISRO, MISR and ASTER from NASA etc. on spacecrafts). The multi-line array sensors acquire images with along or across track stereo viewing with respect to the flight direction. As each line of the image is acquired independently with different exterior orientation by these sensors, bundle adjustment process is unrealistic because of the huge number of unknowns. A piecewise polynomial function depending on time for trajectory modeling can be used to solve such a system [29]. If the GPS/INS (Inertial Navigation Unit) data is provided

still the data does not reflect the physical orientation of the perspective center of the sensor. Hence the required offset vectors and misalignment angles between the GPS, INS and the sensor system have to be measured. Apart from these errors GPS/INS data itself might be corrupted by additional errors due to perturbations, base-station errors etc. For high precision applications these errors have to be modeled and integrated in the bundle adjustment of the image data that eventually results in indirect georeferencing [24] [8].

2.4 Direct Georeferencing

For the sensors carried on aircraft the direct measurement of the external orientation is essential, because of the unpredictable trajectory and the turbulences are not negligible. Although the AT process is considerably improved and expanded to Automated Aerial Triangulation (AAT) in recent years [30], the orientation process using AAT still remains tedious because of large amount of editing and supervision of highly skilled operators. Also, high computational efforts are required to carry out the automatic tie point determination. As a requirement for finding the orientation of each image correctly reliable tie point matching is necessary. This casts a doubt on proper georeferencing of images over an area where there are no obvious features acting as tie points, e.g. large ocean area or hilly terrain.

With the advent of integrated GPS/INS systems the direct georeferencing of the imaging sensors becomes feasible. Specially for the airborne multi-line digital sensors integrated GPS/INS system is the only way to effectively measure the orientation of the sensors. Over the past few years several research efforts have been put into the field of direct georeferencing [32] [31]. The fundamental difference between the direct and indirect approach is the direct measurement of the orientation parameters. The GPS and INS data is optimally integrated using recursive Kalman filtering and the orientation parameters are determined with very high accuracy [32] [31].

If the external orientation parameters are measured using direct georeferencing, a mathematical model has to be adopted accordingly. Now that the orientation sensors are physically displaced from the imaging sensors, additional correction terms have to be used [32]

as even a small error in the orientation of the imaging sensor would project a considerable error on the ground. Assuming an integrated GPS/INS system in combination with an imaging sensor the physical shifts between inertial system and GPS-antenna on the one hand and the perspective center on the other hand are corrected by lever arms defined in the local aircraft body frame. For each system installation these lever arms have to be determined using conventional terrestrial survey methods. The attitudes provided by the integrated GPS/INS system are related on the inertial body frame coordinate axes. Thus an additional misalignment matrix has to be taken into account to transfer the measured attitudes to the imaging sensor frame. Since the misalignment angles between INS and camera frame are not directly observable via conventional techniques they have to be determined indirectly in an appropriate calibration procedure. This attitude transfer is quite a demanding task because reference orientations of superior accuracy are necessary for precise alignment. Although traditional AT provides independent attitude information with high theoretical accuracy, the estimated values are affected by remaining systematic errors and do not agree with the true physical orientation. Photogrammetry provides the only method for determining the misalignment angles in a kinematic airborne environment and the attitude differences between the exterior orientations estimated from AT and GPS/INS at, preferably, several camera stations for the misorientation calibration orientation. The quality of misalignment calibration is strongly dependent on the budget of non modeled systematic errors in the bundle adjustment. The calibrated misalignment angles should remain constant as far as there are no relative movements between the two sensor components. After correcting the GPS/INS orientations by the translational offsets and the misalignment angles, the reduced orientations are interpolated on the exposure times of the imaging sensor to overcome the time offset between the different sensors.

CHAPTER 3

GEORECTIFICATION, IMAGE BACK-PROJECTION AND LIDAR IMAGERY

We present the principles of the georectification process using image back-projection and Lidar imagery. The approach uses concepts of image back-projection. Camera calibration is carried out by using a camera model very similar to that used by Heikkila and Silven [14] [15] to find the intrinsic parameters of the camera. This internal camera model is described as a four-step camera calibration procedure with implicit image correction [15]. This calibration is carried out in the laboratory environment using a well known sized grid. Camera calibration is carried out for the same setting of the camera as it is flown on a mission, thus keeping the camera parameters such as effective focal length and zoom to the original settings. The camera location with respect to a coordinate system is found using the GPS information and the camera pitch, roll and yaw using the kinematic inertial system. Each ground point within the region is then back-projected on the image. Image is then gridded using the nearest neighbor technique. Before projecting the ground points on the image, camera location and the camera view angles (pitch, roll and yaw) are fine tuned using ground control points. The proposed approach uses ground control points to remove the camera location biases and camera view angles to carry out the orthogeorectification. Given a survey flown with the Lidar data, the proposed approach projects the camera imagery after orthogeorectifying the images.

3.1 Motivation

Georectification process requires that each image be georectified using the ground control points contained within the image. This requires a great amount of efforts in terms of manpower in collecting the individual set of ground control points and equally laborious

work thereafter for the process of georectifying individual images. As these images from the digital camera number in thousands, it might not be feasible to process all the images at the same time. This processing speed, however, is at the expense of accuracy. Some loss of accuracy makes up for the gain in speed. If the individual images are being georeferenced each image requires some good 50-60 tie-points to be georeferenced with minimal error. Keeping in mind that each ATM survey collects some 2000 digital camera images and each EAARL survey gathers about 20,000 images, it would become a tedious job georeferencing even the images within a select few regions of interest.

There are applications that do not require as high an accuracy for the georectification. These applications require a broad idea of the location of the feature points on the images. These feature points or objects are viewed to find the location of them on some other kind of imagery such as Lidar DEMs. They just give a general idea of the location of the features while most of the analysis and information extraction is carried out using DEMs. On the other hand some low end applications might just use the georectified digital aerial camera images/mosaics.

3.2 Camera Model with Radial Lens Distortion

Common cameras model the optical system which is far from the ideal optical system. They are often characterized by a distorted projection behavior due to the camera lens distortions. Here we briefly describe the camera model with perspective projection and radial lens distortion. Using this model a point on the ground (w.r.t. world coordinate system) is projected onto a point on the image (w.r.t image coordinate system) through a series of transformations. World coordinates are transformed into camera centered coordinates using the following transformation:

$$\begin{pmatrix} X_k \\ Y_k \\ Z_k \end{pmatrix} = R.\left(\begin{pmatrix} X_w \\ Y_w \\ Z_w \end{pmatrix} + T \right) \tag{3.1}$$

Where R denotes a 3D rotation matrix composed of three partial transformations R_x, R_y, and R_z and T denotes the 3D translation vector given by:

$$T = \begin{pmatrix} T_x \\ T_y \\ T_z \end{pmatrix} = \begin{pmatrix} -X_0 \\ -Y_0 \\ -Z_0 \end{pmatrix}. \tag{3.2}$$

Terms used in equation (3.1) and (3.2) are:
(X_w, Y_w, Z_w) = 3D coordinates of a point P in the world coordinate system,
(X_k, Y_k, Z_k) = 3D coordinates of point P in camera centered coordinate system,
(T_x, T_y, T_z) = Translation vector.

Assuming the normalized pinhole camera projection the transformation of the (X_k, Y_k, Z_k) point in camera coordinate system into the undistorted image coordinate system point (x_u, y_u) can be given using geometrical ray theorems as

$$x_u = \frac{f_k.X_k}{Z_k} \text{ and } y_u = \frac{f_k.Y_k}{Z_k}$$

Where f_k is camera's effective focal length.

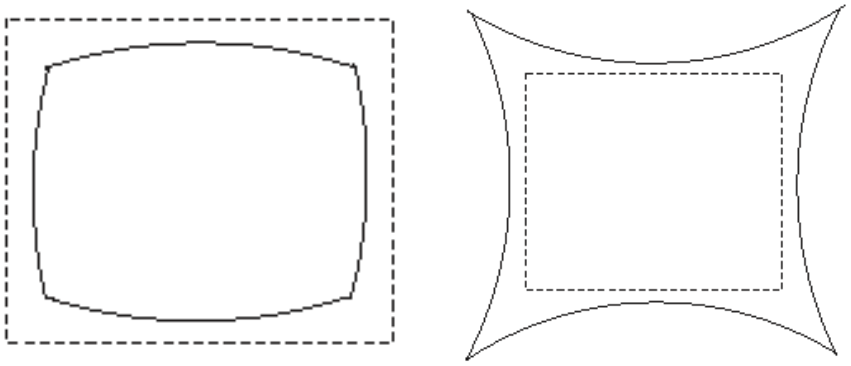
Distorted point coordinates after including the lens distortion using the plumb bob's model [7] are

$$X_d = (1 + \kappa_1.r^2 + \kappa_2.r^4 + \kappa_5.r^6)X_u + D_t \tag{3.3}$$

Where $\kappa_1, \kappa_2, \kappa_5$ are the radial distortion coefficients for the second, fourth and sixth order radial distortion respectively, $r^2 = x^2 + y^2$, and D_t is the tangential distortion vector given by

$$\begin{pmatrix} 2.\kappa_3.x.y + \kappa_4.(r^2 + 2.x^2) \\ \kappa_3.(r^2 + 2.y^2) + 2.\kappa_4.x.y \end{pmatrix} \tag{3.4}$$

κ_3 and κ_4 are the tangential distortion coefficients. Imperfect centering of the lens components and other manufacturing defects in a compound lens are the causes of tangential distortion.



(a) (b)

Figure 3.1 Effect of Radial Distortion on Image Geometry. Dashed Lines Represent the Image as it Would Appear in the Absence of Radial Distortion. Solid Lines Represent the Image Geometry in Presence of (a) Barrel and (b) Pincushion Radial Distortion.

If the augmented vector of X_p is denoted by $\widetilde{X}_t$ that is the vector created by adding 1 as the last element of X_t. The final pixel coordinate vector X_p of the projection of P on the image plane can now be given in terms of its augmented vector as

$$\widetilde{X}_t = A \cdot \widetilde{X}_d \tag{3.5}$$

Where X_d is the augmented distorted point coordinate vector and A is called the camera intrinsic matrix which is defined as

$$A = \begin{pmatrix} f_x & \alpha \cdot f_x & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{pmatrix} \tag{3.6}$$

Matrix elements are explained below.

f_x and f_y are the effective focal length of the camera (a unique value) expressed in terms of horizontal and vertical CCD pixel units respectively. The expression f_y/f_x is called aspect ratio, and is different from 1 if the pixels in the CCD array are not square.

(c_x, c_y) is the principal point of the image i.e. the intersection point of principal axis with the image plane.

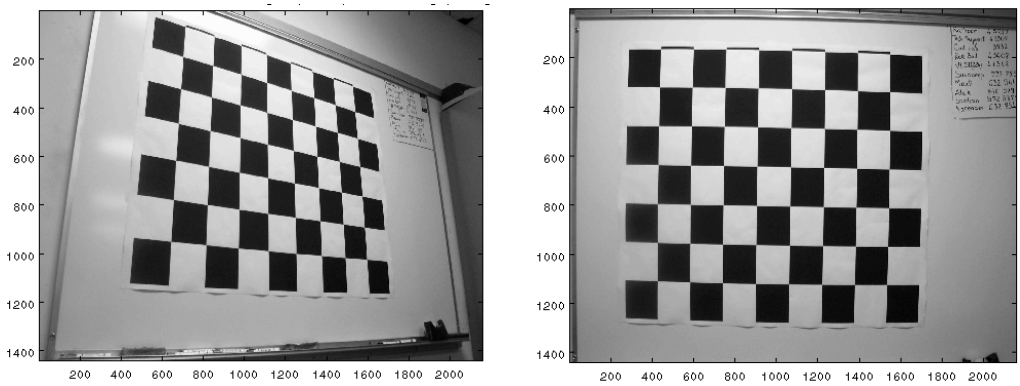
α is the angle between x and y axes of the CCD array.

3.3 Camera Calibration

Camera calibration uses a calibration model that requires multiple images from different orientations for camera calibration. This technique requires a camera to view a planar set of target points from a few orientations. Either the camera or the planar set of points is movable and can be shifted for acquiring different orientation images. This calibration technique does not require the motion of the camera to be known when shifted from one position to the other. This calibration technique models the distortion camera model. Using this process, camera calibration problem is solved by a closed-form solution, followed by a nonlinear optimization technique based on the maximum likelihood criterion. Finally, lens distortion is taken into account giving analytical and nonlinear solutions. Camera calibration is carried out using matlab toolbox.

A well defined square grid (as shown in Figure 3.2) is printed using precision plotters and is placed on a plane surface. This technique requires a minumum of two noncoplanar images for the calibration, if the skewness constraint is assumed to be zero. If the skewness constraint is included then at least three images are required to carry out the calibration, more the number of images better is the estimation of the parameters. Intrinsic and extrinsic parameters of the camera are calibrated.

Other essential paramters such as the rotation matrix and the translation vector are extracted using the Inertial Navigation System (INS) and GPS systems respectively. The INS provides the pitch, roll and the yaw angles of the aircraft which can be used to find



(a) (b)

Figure 3.2 Two Non-Coplanar Set of Images for Calibration, (b) Also Shows the Visible Lens Distortion at the Corners.

the orientation of the camera w.r.t. some known coordinate system. Rotation matrix for a single image camera position is calculated as

$$R = BCD \tag{3.7}$$

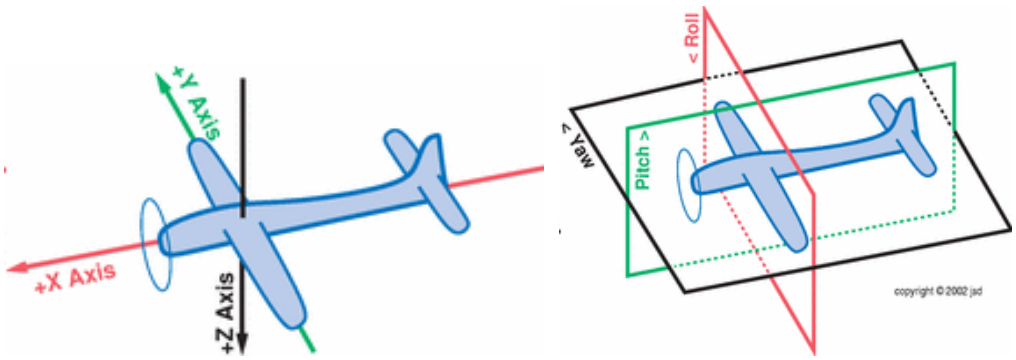
Where,

$$D \equiv \begin{pmatrix} \cos\phi & \sin\phi & 0 \\ -\sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{3.8}$$

$$C \equiv \begin{pmatrix} \cos\theta & 0 & -\sin\theta \\ 0 & 1 & 0 \\ \sin\theta & 0 & \cos\theta \end{pmatrix} \tag{3.9}$$

$$B \equiv \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\psi & \sin\psi \\ 0 & -\sin\psi & \cos\psi \end{pmatrix} \tag{3.10}$$

In this convention, θ is pitch, ϕ is yaw, and ψ is roll (Figure 3.3).



(a) (b)

Figure 3.3 (a) X, Y and Z Axes Represented with the Aircraft and (b) Pitch, Roll and Yaw Rotations of an Aircraft i.e. the Camera.

Next the camera location is established with respect to some known coordinates. This is done by projecting the flight track information, which is captured in units of degrees along latitudes and longitudes, using the Universal Transverse Mercator (UTM) projection. Camera or the flight track points are projected on a 2D grid as Eastings and Northings in meters, which in turn represent orthogonal X-Y axes on the ground. Although the Earth's surface is not flat, for small regions it's assumed to be flat.

Digital Elevation Model (DEM) for the to-be-mapped region is extracted using the LIDAR survey flown the very same day as the digital camera images, so as to minimize the back-projection errors. The LIDAR data is feeded through a processing system, called LaserMap, and the output at the final level is a DEM for the associated region. This data is appropriately gridded to the required resolution using the standard gridding by averaging a 3x3 kernel. The camera/aircraft height as well as the DEMs are measured with respect to AGL Average Ground Level Ellipsoid (AGL) in meters.

Corrections are made to the time-stamping of the images as well as the mounting biases of the camera. These corrections are made by using well known ground control points, also known as tie-points. Location of these points corresponding to some well known ground points is well established on the image plane. Best estimates for the time-stamping and mounting errors can be obtained by minimizing the residual between the back-projection through the model and the N observations (U_i, V_i), where i = 1,...,N. Thus the objective function is expressed as a sum of squared residuals:

$$F = \sum_{i=1}^{N}(U_i - u_i)^2 + \sum_{i=1}^{N}(V_i - v_i)^2 \tag{3.11}$$

The least squared estimation technique can be used to minimize Eq. 3.11. Due to the nonlinear nature of the camera model, simultaneous estimation of the parameters involves applying an iterative algorithm. For this problem the Levenberg-Marquardt optimization method has been shown to provide the fastest convergence. However, without proper initial parameter values the optimization may stick in a local minimum and thereby cause the estimation to fail. This problem can be avoided by using the parameters from the Direct Linear Transform (DLT) method as the initial values for the optimization. A global minimum of Eq. 3.11 is then usually achieved after 8-10 iterations.

It should be noted that the mounting bias errors and the time-stamp bias are constant throughout the survey as the camera mounting base is fixed and also the camera captures images in a fixed time interval. The situation, however, changes for different surveys as the time-stamping errors might be entirely different and if the camera is unmounted during the next survey mounting bias is most likely to change when the camera is remounted. As the aircraft speed is at about 50 meters/second, even a fraction of difference in the time-stamping will create a huge error on the ground. Similarly, fraction of error in estimating the mounting bias would create a huge error on the ground depending on the altitude of the aircraft. The higher the altitude of the aircraft more is the effect of mounting biases on ground.

These errors (i.e. the mounting biase error and the time-stamping error) are now incorporated into the back-projection technique and the particular region points are back projected onto the image. These positions give the actual radiation emitted by the ground points recorded on the image. These back-projected points are gridded using the nearest neighbor technique to fit onto the output image.

The results mostly depend on the calibration technique used. The four-step procedure for camera calibration can be utilized in various machine vision applications, but it is most beneficial in camera based 3-D measurements and in robot vision, where high geometrical accuracy is needed. This procedure uses explicit calibration methods for mapping 3-D coordinates to image coordinates and an implicit approach for image correction. Experiments have shown that the standard error in fitting residuals using four-step camera calibration approach, under realistic conditions would always be less than 0.01 pixels and the feature detection accuracy would be about 0.02 pixels. For more intensive distortions (as is in our case) the errors are slightly bigger [15].

CHAPTER 4

EXPERIMENTAL RESULTS

We present detailed experimental results with the discussion about the software processing system that supports it. We compare the results with another approach used by us to georectify the digital camera images.

4.1 Camera Calibration

Camera used along with the Lidar surveys in the ATM3 instrument is a Kodak DC4800. This camera is adjusted at full zoom out for the surveys and hence it is easier to calibrate. The camera is set to take images every ten seconds. This time interval is set depending upon the altitude of the aircraft, aircraft's speed and the overlap between the successive images. This in turn insures that the required area is covered by the camera photograph.

With camera put to the same settings as it is flown on the surveys this ensures the accuracy of the georectification. There is another reason as why not carry out the calibration for the camera using the 3D information of the ground control points. We chose to use Heikkila's calibration as it finds the radial distortion coefficients up to sixth order and also finds tangential distortion coefficients.

A well known, chessboard-like, square grid is used as a target for the camera calibration (Figure 3.2). This grid is printed using high resolution plotters and grid-dimensions are well known. Each square within this grid measures 125.833mm in the X-direction and 124.686mm in the Y-direction. Six to seven noncoplanar images are captured using exactly the same settings for the camera as used in the surveys. Minimum requirement for the number of images is three, but for the algorithm to converge nicely and also for the minimum estimation errors more images can be taken. Images are taken from a distance sufficiently far from the grid so as to cover the whole grid at the same time near enough so

as to have a nice distribution of the target points. Corners of this grid are extracted with sub-pixel accuracy and provided to the algorithm to find the camera's intrinsic and extrinsic parameters. The purpose of this calibration is to get the camera's intrinsic parameters. Three different target grids with different spacing of the target points on the grid are used for the calibration to assess the parameter values. Results of the camera calibration are shown in the table 4.1.

Table 4.1 Camera Intrinsic Parameters.

Parameter	Description	Value	Error
f_x	Focal Length (Units of horizontal CCD pixels)	1740.39960	5.93592
f_y	Focal length (Units of vertical CCD pixels)	1744.08383	5.44191
c_x	Principal Point on horizontal axis	1077.77280	6.72177
c_y	Principal Point on vertical axis	693.29595	5.22348
α	Skew Coefficient	0.00	0.00
κ_1	First order distortion coefficient	-0.11102	0.00898
κ_2	Second order distortion coefficient	0.04104	0.02241
κ_3	Third order distortion coefficient	0.00165	0.00080
κ_4	Fourth order distortion coefficient	-0.00035	0.00098
κ_5	Fifth order distortion coefficient	0.00	0.00

4.2 Georectification

We present the detailed experimental results for the georectification. We also discuss an approach not requiring the camera calibration and image correction, due to which it is faster but is much less accurate than the proposed approach. For applications requiring a bird's-eye-view of feature point locations, this approach, we call it "image pasting", could give an idea of where desired features are located. But for a more intense quantitative and qualitative analysis this approach lacks the desired accuracy. Furthermore, this approach does not preserve the feature shapes.

4.2.1 Image Pasting

Before we begin showing the results for the proposed method, we would like to explain in brief another method to carry out georectification and would also like to discuss errors

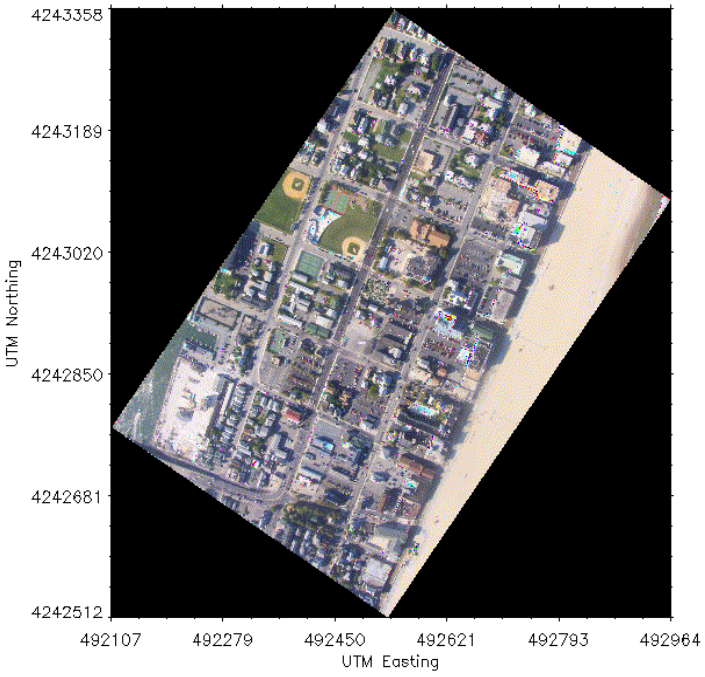


Figure 4.1 A Georectified Image over Ocean City, Maryland.

involved using this method. In this method we extract the camera's position, altitude and corresponding pitch, roll and yaw values for the aircraft (i.e. the camera's rotation angles). Camera's position in latitude-longitude is projected into Universal Transverse Mercator (UTM) coordinates. Camera's view angles are calibrated by placing a grid at a distance and capturing the image placing the camera perfectly perpendicular to the grid. The view angles of camera in image's X and Y directions are given by

$$\alpha_x = 2 \cdot \arctan\left(\frac{Distancecapturedin X - direction}{2 \cdot Distancebetweencameraandthegrid}\right)$$

Similarly,

$$\alpha_y = 2 \cdot \arctan\left(\frac{Distancecapturedin Y - direction}{2 \cdot Distancebetweencameraandthegrid}\right)$$

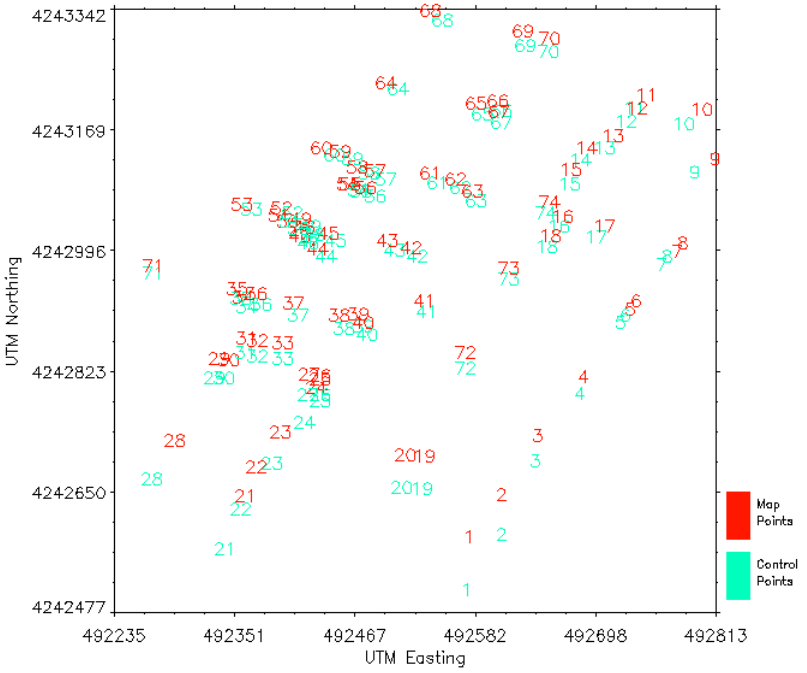


Figure 4.2 Comparing Target Points on the Map with Ground Control Points.

Due to pitch and roll of the aircraft image's center point is not the point where the aircraft/camera is and hence the image center point needs to be shifted for this discrepancy.

After ascertaining the images center point on the ground image is linearly stretched for the area that it covers on the ground. This image is rotated and gridded according to the output map resolution. One such map using this technique is shown in figure 4.1.

The ground control points are laid down on the georectified image (see Figure 4.2). As we can see form the geocorrected image that these control points have a lot of error.

Table 4.2 shows the actual ground control point locations and the map point locations after geocorrecting the image by using the above described method.

Table 4.2 Control point and map point locations.

Control Point#	Actual Northing	Actual Easting	Map Northing	Map Easting
1	4242502.059	492569.728	4242577.84	492571.72
2	4242581.156	492602.727	4242637.08	492603.24

Table 4.2 continued

Control Point#	Actual Northing	Actual Easting	Map Northing	Map Easting
3	4242686.160	492634.820	4242721.81	492636.99
4	4242782.889	492677.182	4242807.58	492681.25
5	4242885.047	492716.642	4242903.33	492726.03
6	4242895.542	492721.681	4242915.04	492731.36
7	4242968.170	492756.398	4242986.85	492771.15
8	4242980.211	492761.702	4242997.70	492776.66
9	4243100.301	492788.786	4243119.63	492808.34
10	4243168.835	492772.033	4243190.58	492789.23
11	4243191.175	492724.829	4243210.56	492736.19
12	4243172.236	492716.321	4243191.44	492727.40
13	4243134.510	492696.419	4243152.52	492703.64
14	4243118.455	492672.790	4243135.47	492678.49
15	4243082.471	492662.672	4243104.47	492663.86
16	4243021.897	492652.324	4243036.97	492656.28
17	4243006.410	492687.804	4243022.85	492696.40
18	4242993.974	492640.820	4243008.90	492644.05
19	4242647.075	492521.150	4242693.74	492522.98
20	4242648.596	492501.699	4242695.29	492504.04
21	4242560.682	492331.160	4242635.19	492351.63
22	4242616.759	492346.913	4242676.87	492361.79
23	4242682.992	492376.770	4242728.19	492385.21
24	4242740.606	492407.293	4242791.37	492419.09
25	4242772.916	492422.939	4242803.79	492422.93
26	4242780.497	492422.167	4242808.61	492423.27
27	4242781.014	492411.138	4242809.47	492412.42

Table 4.2 continued

Control Point#	Actual Northing	Actual Easting	Map Northing	Map Easting
29	4242660.234	492260.821	4242715.96	492283.26
30	4242805.237	492321.344	4242832.55	492325.97
31	4242805.264	492330.862	4242831.34	492335.44
32	4242840.844	492351.204	4242862.69	492351.98
33	4242835.390	492362.761	4242857.86	492362.82
34	4242832.852	492386.629	4242854.94	492386.42
35	4242906.114	492352.748	4242921.58	492349.39
36	4242919.296	492346.555	4242932.95	492343.36
37	4242911.195	492365.619	4242925.89	492361.79
38	4242894.993	492402.158	4242911.42	492397.44
39	4242876.202	492445.299	4242894.55	492441.53
40	4242879.981	492463.965	4242896.96	492459.61
41	4242866.839	492467.779	4242884.73	492464.26
42	4242900.237	492525.299	4242916.24	492522.81
43	4242978.762	492515.826	4242991.50	492510.41
44	4242988.733	492493.784	4243001.66	492486.82
45	4242978.970	492429.009	4242990.12	492420.86
46	4243001.691	492438.132	4243013.03	492430.51
47	4242998.271	492411.385	4243008.38	492402.61
48	4243001.526	492413.101	4243011.65	492403.47
49	4243010.878	492417.067	4243021.47	492407.94
50	4243022.214	492413.703	4243032.83	492404.33
51	4243020.018	492401.453	4243030.21	492391.69
52	4243030.242	492394.142	4243039.03	492384.01
53	4243040.784	492395.882	4243049.37	492385.38

Table 4.2 continued

Control Point#	Actual Northing	Actual Easting	Map Northing	Map Easting
54	4243047.291	492356.718	4243052.98	492347.50
55	4243074.175	492459.474	4243083.64	492448.76
56	4243072.356	492462.106	4243082.95	492451.69
57	4243065.676	492475.952	4243077.61	492466.67
58	4243089.387	492485.790	4243101.38	492476.14
59	4243097.754	492470.016	4243107.48	492458.43
60	4243119.755	492454.360	4243128.93	492442.73
61	4243124.444	492436.177	4243135.47	492423.96
62	4243084.541	492535.754	4243098.28	492528.84
63	4243077.145	492557.759	4243089.66	492553.47
64	4243059.383	492572.365	4243072.62	492569.31
65	4243219.183	492497.511	4243228.81	492485.96
66	4243183.912	492577.884	4243198.85	492573.10
67	4243188.063	492596.856	4243203.01	492593.65
68	4243171.026	492596.540	4243186.10	492594.28
69	4243317.059	492540.655	4243331.38	492529.01
70	4243281.534	492619.755	4243301.49	492617.87
71	4243272.458	492641.944	4243291.67	492642.67
72	4242954.912	492262.755	4242965.50	492262.25
73	4242819.549	492562.344	4242842.02	492562.42
74	4242946.614	492603.827	4242962.40	492604.10
75	4243041.430	492639.445	4243056.43	492642.85

37 more ground control points were taken from a different location in Assateague Island. The image is again geocorrected using the above described method (Figure 4.3). The actual

ground point locations and the map point locations for the same points using the above method is listed in the table 4.3.

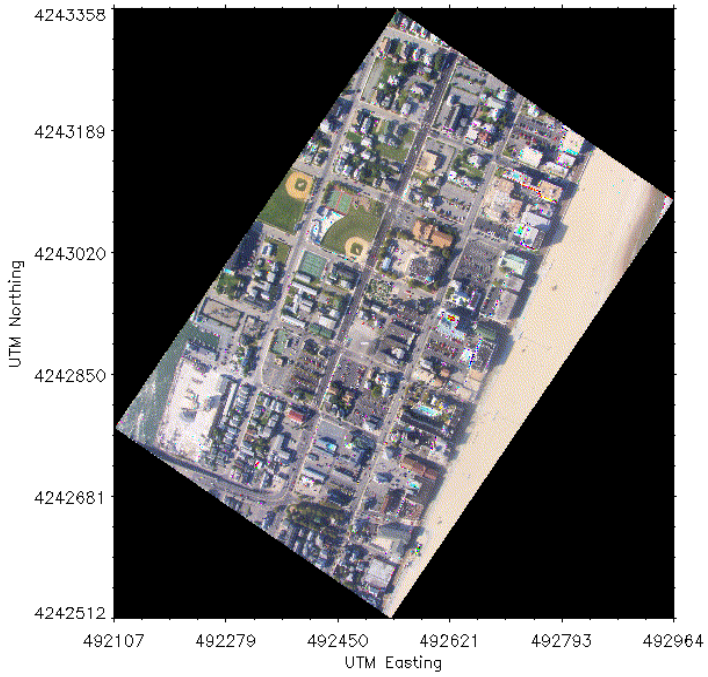


Figure 4.3 A Georectified Image over Assateague Island, Maryland.

4.2.2 Image Back-Projection

The proposed approach is used to georectify the images containing the ground control points. The image ground control points are still in error before using the bias correction. We show the results before the bias correction first and then we will show the results after using the bias correction.

The ground control points are back-projected using the camera model described in section 3.2. Camera intrinsic parameters are extracted using the camera calibration. Camera intrinsic parameters remain unchanged in both the cases viz. before bias correction and after bias correction. Bias correction includes camera mounting bias for pitch, roll and yaw and image time-stamping. Before using the bias correction these parameters are extracted as provided by the Inertial Navigation System (INS) and the Global Positioning System (GPS).

Figure 4.4 and figure 4.5 depict the georectified image without the bias correction. Table 4.2 compares the actual control point locations with the locations obtained using this approach without bias correction.

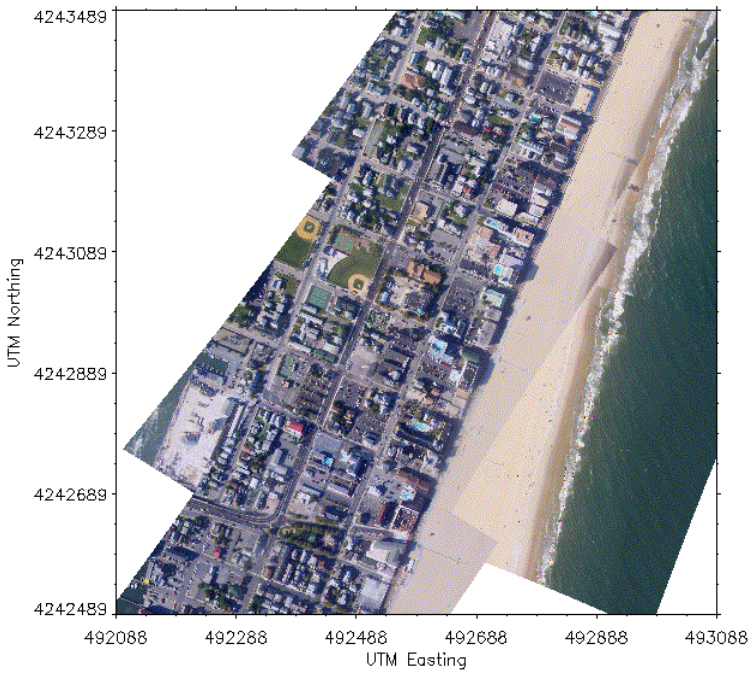


Figure 4.4 A Georectified Image over Ocean City, Maryland.

As can be inferred from the table 4.2 that location errors are still large as we have not yet taken mounting bias and the time-stamping errors into account. Table 4.4 shows the results for the mounting and time-stamping bias and its affect on ground.

After incorporating the bias values in the back-projection approach georectified images gave ground control point locations very close to the actual values. These results are shown in the Figure 4.6, Figure 4.7 and Table 4.4. In case of preparing digital camera image mosaics the consecutive images are band equalized. This is done as follows:

- Find the common projected points within two overlapping images.
- Individually compute the mean values for all the bands for both the images.
- Compute the difference in the mean band value for each band.

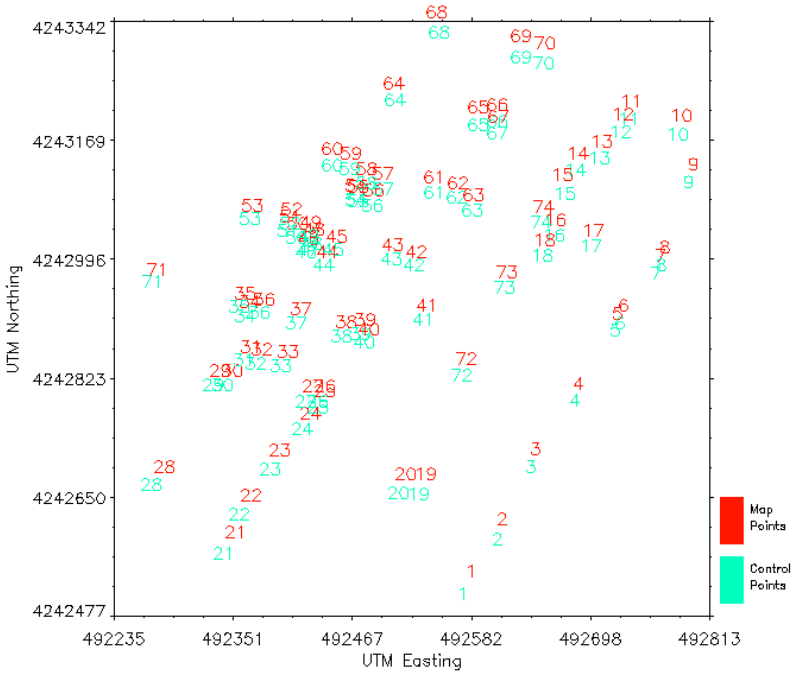


Figure 4.5 A Georectified Image over Assateague Island, Maryland.

- Assuming we are equalizing the mosaic with respect to first image to be projected on the region, we compensate the band values in the second image with the differences found in the previous step.

- Continue this process for each image, equalizing it with the previous image pasted until the last image points in the region are projected.

The above stated process works whenever overlapping projected points in both the images (i.e. the image to be equalized and the equalized image) represent the same point on the ground, hence implying better quality of georectification for a better image equalization. Image equalization quality also depends on the number of overlapping projected points, more the number of such common points better is the representation of reflectance of the ground points at the time of capture of two images.

It was observed that the minimum mean location error was found using the image back-projection after bias removal method. The results with three different methods are compared in the figures 4.8 and 4.9 for two different sets of ground control points.

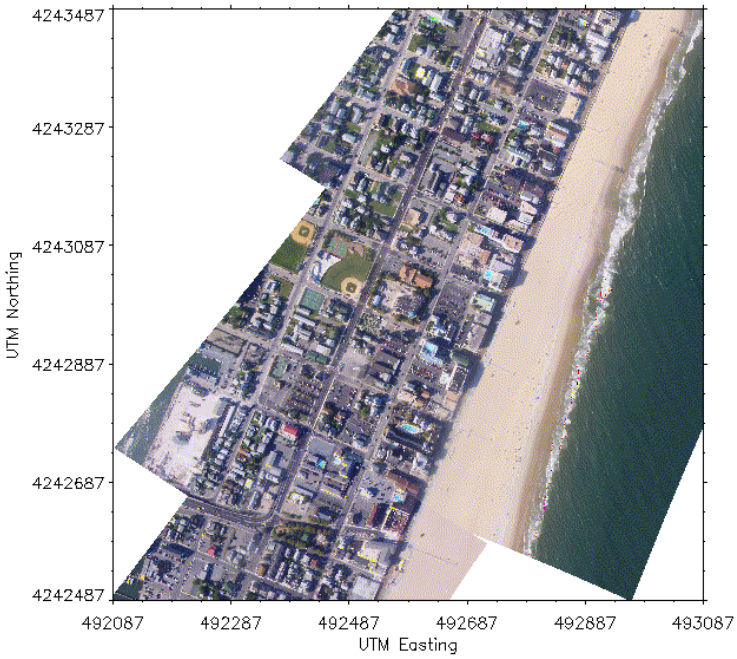


Figure 4.6 A Georectified Image over Ocean City, Maryland.

4.3 Park-View Processing System

With every mission flown with ATM3 instrument approximately 2000 digital camera images are acquired along with a large amount of LIDAR data, covering a large land area along the coast. For intensive research analysis of coastal topographic change and information extraction LIDAR data is used but some applications require the georectifed digital aerial photos. There are several other types of imageries that are used for extracting information but aerial photography has its own place. Coastal regions where rate of change is much higher georectified digital photography can be used to determine those changes and analyzing the impact of storms or hurricanes or even in warning out before such natural calamities take place. Usually park managers require such kind of means to have a quick-look-type maps. Park-view is based on one such concept which provides this ability to view a quick-look map of any desired region of the surveyed area.

The 'Park-View' main Graphical User interface (GUI) (see Figure 4.10) is built using IDL-GUI toolkit. The fields used in this GUI are:

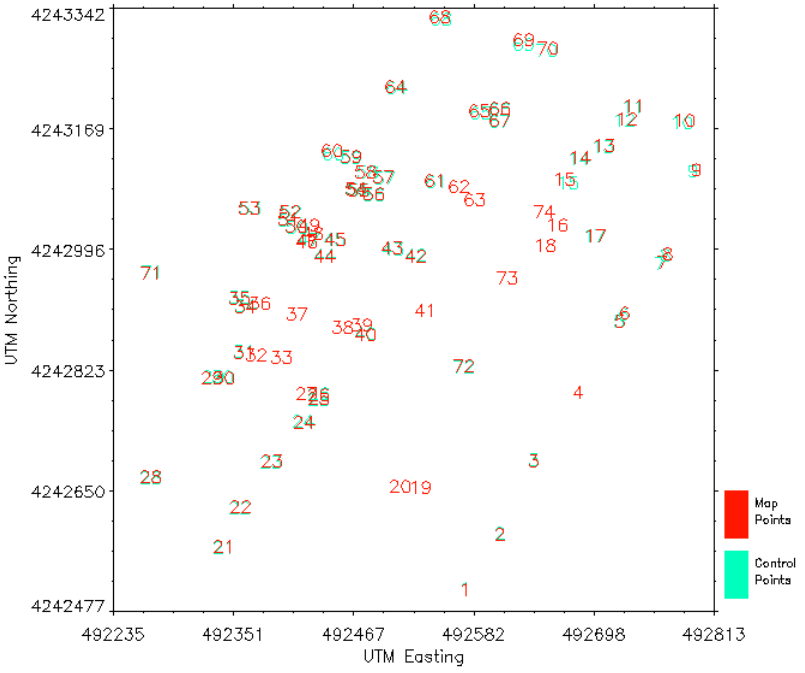


Figure 4.7 Projected Points Compared to Actual Ground Control Points. With the Biases.

- Survey date
- Input path of the data
- Coastline file name
- Option for having the flight track in UTM coordinates
- Type of instrument used. As of now there are two different types of instruments carrying two different types of cameras - ATM3 carries a Kodak DC4800 with 3 bands (RGB) and an image resolution 2160x1440 pixels and EAARL instrument carries AXIS 2100 with 3 bands and an image resolution 352x240 pixels.
- Desired region coverage in mxm.
- Optional field for providing the latitude
- Optional field for providing the longitude

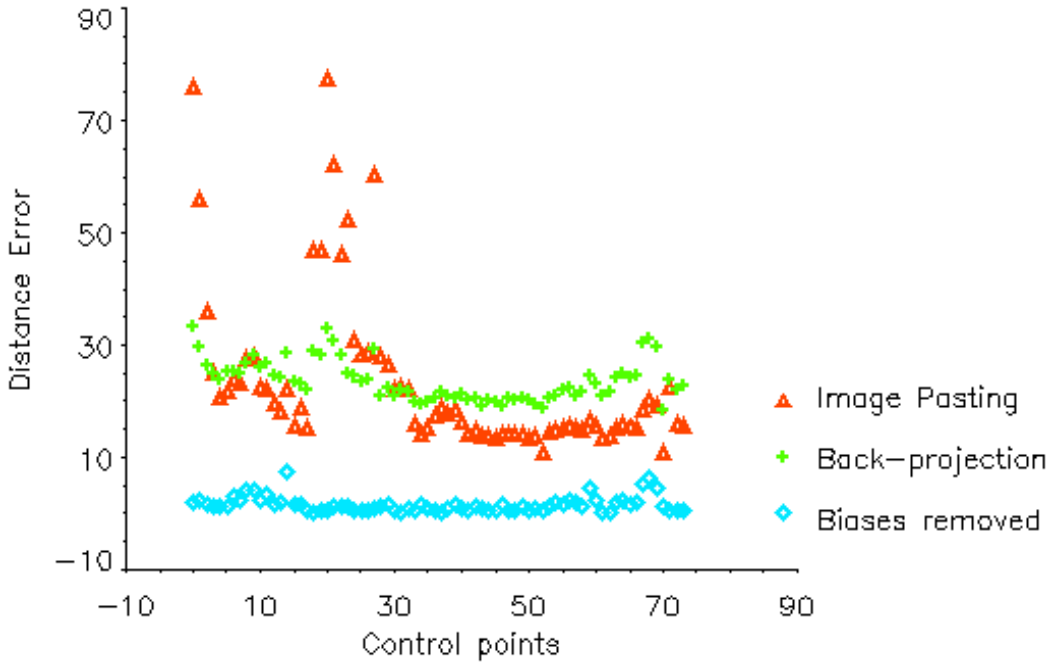


Figure 4.8 Distance Errors Corresponding to Different Methods over a Set of 74 Control Points.

The main GUI provides the facility to map a coastline using the 'coastline file' and map the flight track over the coastline. These maps are displayed on the window within the main GUI. User has an option of choosing the type of coordinates – latitude-longitude or Universal Transverse Mercator (UTM) – for the display. User can select an area from the displayed map and provide width of the region to be displayed with georectified digital camera images/mosaics. Once the process is started georectified image of the region is displayed (Figure 4.11). The georectified image can also be saved as a JPEG map of the region or as a GEOTIFF which is widely used and is supported by many Geographic Information System (GIS) softwares. Location of a point in the displayed window can be ascertained by placing the mouse pointer at that point, this location is displayed in UTM coordinates only.

The mapper GUI for the 'Park-View' (Figure 4.12) has the option of georectifying all the images individually and store them as JPEG maps or GEOTIFFs. Fields used in this GUI are pretty much the same as used in the main GUI.

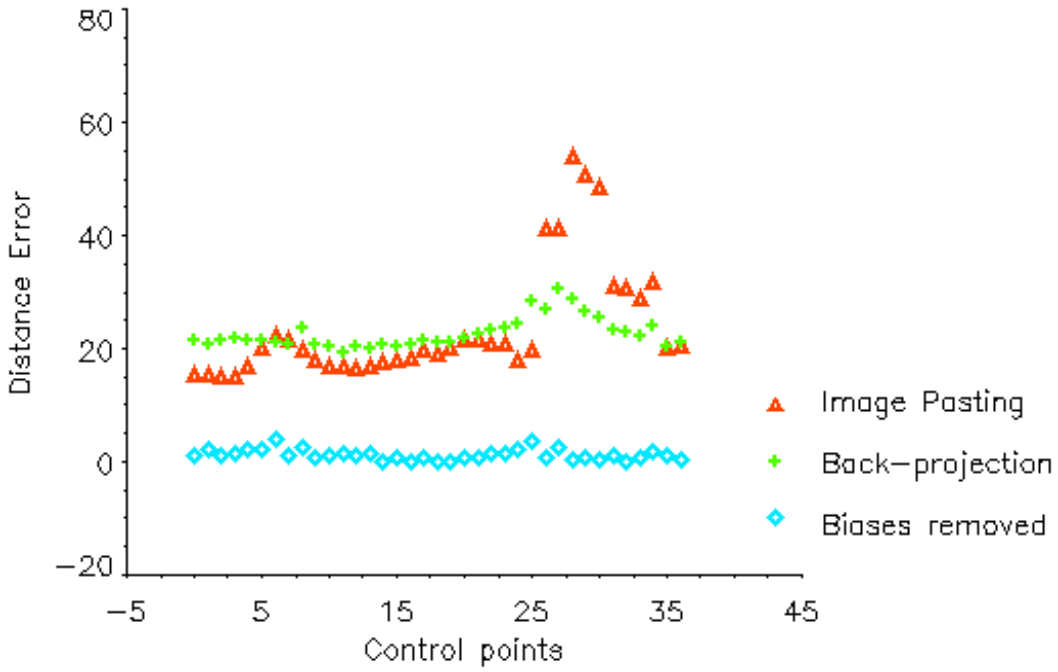


Figure 4.9 Distance Errors Corresponding to Different Methods over a Set of 37 Control Points.

The Park-View processing system is faster in terms of processing time for all the images combined in any given survey. It also provides the option to georeference individual images captured in the survey separately, which could be done easily once the camera calibration is carried out and the errors in mounting and time-stamping have been quantified. Map mosaics for any region of interest can also be prepared for any given size. Because of using polynomial transformations to model the camera behavior, this approach performs much better in terms of accuracy. Still there is a lot of scope for improving accuracy of the georeferenced images and mosaics. Georeferencing accuracy eventually depends on the accuracy of camera calibration and the accuracy of tie-points (also called ground control points), which in turn depends on the ability to match distinguished feature points on the images. There are other factors contributing towards error but are not as prominent as the above stated ones, this includes errors with GPS and INS systems and errors involving improper focusing of the images.

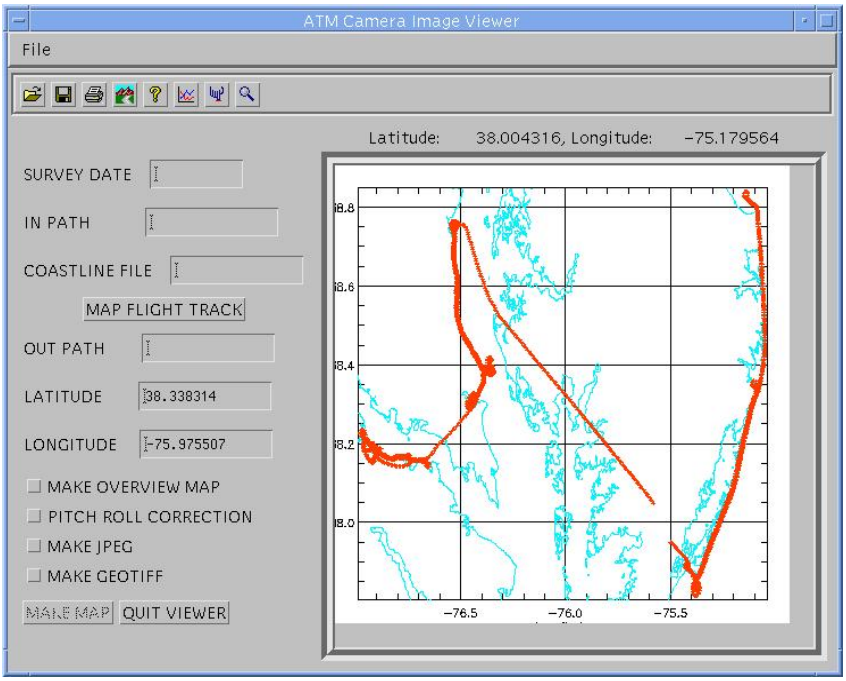


Figure 4.10 Park-View Viewer GUI.

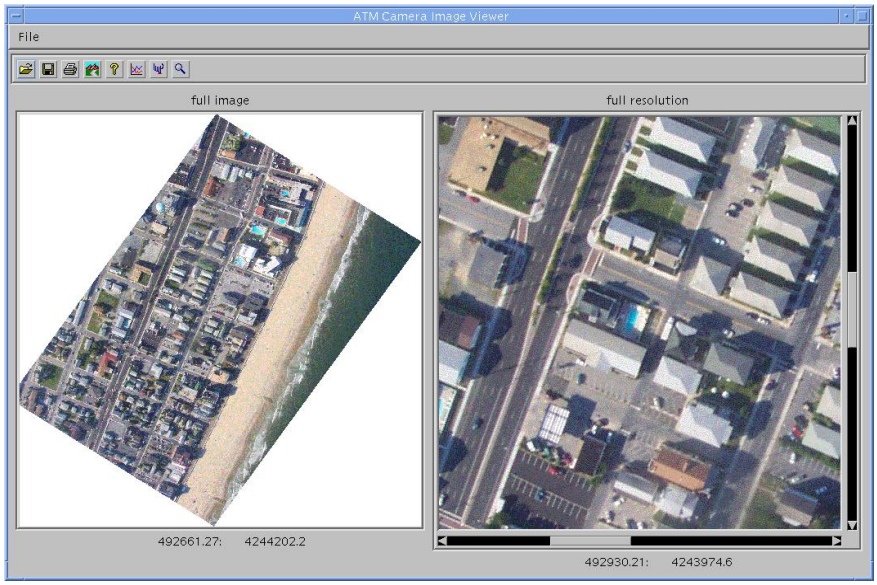


Figure 4.11 Park-View Map Display Window.

Table 4.3 Control Point and Map Point Locations.

Control Point Point#	Actual Northing (m)	Actual Easting (m)	Map Northing (m)	Map Easting (m)
1	4228788.592	486715.319	4228801.21	486706.65
2	4228765.845	486707.602	4228778.16	486698.27
3	4228816.186	486724.848	4228827.74	486715.55
4	4228768.603	486740.278	4228781.48	486732.84
5	4228718.309	486722.120	4228733.30	486715.03
6	4228667.943	486704.893	4228686.68	486697.57
7	4228619.404	486693.715	4228641.12	486689.71
8	4228663.457	486587.533	4228683.02	486578.68
9	4228691.253	486596.156	4228708.33	486586.71
10	4228740.706	486613.248	4228754.94	486602.43
11	4228764.451	486621.681	4228775.72	486609.23
12	4228792.474	486631.335	4228804.00	486619.19
13	4228815.138	486639.514	4228824.95	486626.34
14	4228842.939	486649.667	4228852.53	486636.12
15	4228845.762	486617.528	4228854.45	486602.25
16	4228856.852	486621.077	4228865.45	486605.39
17	4228854.182	486605.138	4228862.83	486589.33
18	4228894.268	486596.494	4228902.29	486578.51
19	4228914.618	486601.469	4228921.67	486583.92
20	4228927.813	486595.099	4228934.58	486576.41
21	4228953.697	486581.551	4228959.55	486561.05
22	4228984.488	486596.748	4228990.62	486576.24
23	4228970.037	486609.295	4228977.88	486590.03
24	4228982.075	486627.610	4228990.45	486608.71
25	4229007.519	486676.397	4229017.51	486661.61
26	4229089.674	486710.242	4229102.35	486695.30
27	4228512.811	486635.964	4228553.66	486633.15
28	4228471.599	486614.154	4228511.41	486624.42
29	4228451.047	486594.545	4228504.95	486597.19
30	4228476.974	486489.901	4228526.95	486498.20
31	4228489.962	486494.394	4228537.77	486501.87
32	4228576.087	486523.941	4228607.08	486523.34
33	4228588.202	486528.706	4228618.60	486526.83
34	4228593.047	486521.782	4228621.74	486520.03
35	4228585.11	486511.918	4228616.68	486511.47
36	4228669.319	486556.237	4228687.91	486548.48
37	4228682.239	486561.027	4228700.82	486553.02

Table 4.4 Mounting Bias and Time-stamping Bias.

Pitch Bias (degrees)	Roll Bias (degrees)	Yaw Bias (degrees)	Time-stamp Bias (seconds)
1.4999995	1.0000000	0.80000305	0.0350225

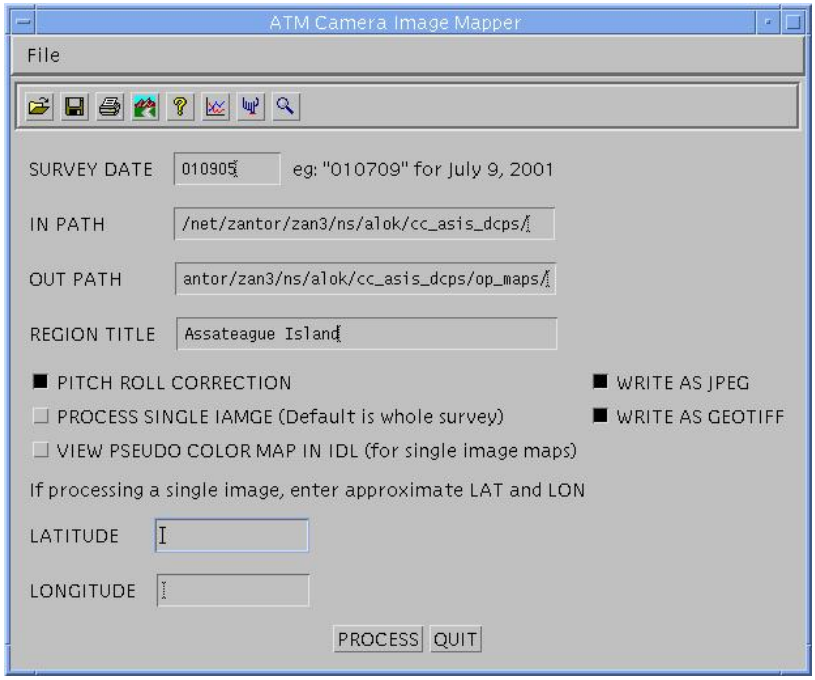


Figure 4.12 Park-View Mapper GUI.

CHAPTER 5

CONCLUSIONS

Park-View is a processing system for georeferencing the digital camera images captured by Kodak DC4800 camera on-board NASA Airborne Topographic Mapper (ATM) apart from the LIDAR data collected for the topographical analysis. ATM has been used to survey much of the US coastline topography as part of the NASA/NOAA/USGS funded mapping project. The inclusion of the ability to capture 3-band or multispectral 0images along with LIDAR data has certainly enhanced the capability to carry out scientific analyses. Furthermore it provides the required data set for the analyses in much shorter time than the traditional techniques where georeferencing has to be done manually and also all the different images have to have the Ground Control Points (GCPs). Collecting of ground control points with each survey is a difficult task as with each survey on an average 2000 images are captured and for each image to be georeferenced manually it needs to have at least three GCPs. This, however, comes in expense of the accuracy of georeferencing. So, for the applications not requiring very high accuracy Park-View provides with the georeferenced images and mosaics in almost no time. Functioning of the Park-View processing system is summarized below:

- The ATM collects data using LIDAR as spot elevation for each point on the surface it strikes at. This data is calibrated at Level1 processing to give a 14 word information about each spot elevation. This data consists of the description about point location and its topography. This data is processed further by passing through the different levels of the LaserMap processing system to get the Digital Elevation Model (DEM). The final DEM contains only the point locations with some reference coordinate system and the spot elevation with reference to some datum.

- Digital camera is calibrated in the lab for extracting its intrinsic parameters, which are later used in camera model to georeference the digital camera images. Well defined square grids are used to capture the calibration images. Multiple images are taken from all different directions of a single grid. Target points are a planar set of points on the grid image. These images are then calibrated using Heikkila's calibration model. Camera intrinsic parameters such as focal length, principal point, skew coefficient and radial distortion coefficients are extracted.

- Corrections are made to the location of the aircraft, in effect to the camera, and its attitude. Correspondences are found for the GCPs in the images and the location and attitude parameters are adjusted in least squared manner to get the exact location and attitude for the given image. Differences in the least error attitudes and the attitude provided as per the time-stamp of the image are noted. The process is applied to get the location differences and accordingly the time-stamp error is found. These attitude and time-stamp errors should be the same for all the images taken in a survey as the camera remains fixed on the same mount for each survey. Hence, for rectifying other images these corrections could be applied.

- Finally the images are georeferenced using the camera model with radial lens distortion. The required parameters and the data is acquired in the above stated first three steps. Step three overcomes the errors caused by the mounting biases and the time-stamping error.

  Park-View provides the ability to process the images individually by georeferencing each image in a batch process. It also provides the ability to prepare the mosaics of a region specified by the user. In case of mosaics, the images are color equalized to make them appear as one big smooth image. Efforts are underway to analyze these georeferenced images and extract useful information using Park-View.

5.1 Future Research

The processing system for Park-View would give an additional advantage in analyzing topographic data collected with NASA's Airborne Topographic Mapper (ATM). This processing system could be easily incorporated into georeferencing digital camera images, acquired coupled with other instruments such as Experimental Airborne Advanced Research LIDAR (EAARL). The information extracted, using topographic data, could be laid over georeferenced images or mosaics and hence could provide a better understanding for the interpretational and analysis purpose.

A couple of improvements suggested for the enhancement of the processing system – Park-View – are as follows:

- Camera calibration is done using Heikkila's method which uses noncoplanar set of images for calibration. This introduces some errors as the calibration grid image may not exactly be placed in a perfect plane. This method however has an advantage of extracting distortion coefficient to the fifth order of the radial and tangential distortions. Advantage of using Tsai's model is that it uses a 3D grid for the calibration, which is more accurate than the 2D noncoplanar set of images. While disadvantage of using Tsai's model is that it only approximates the first order radial distortion. The digital photography camera (Kodak DC4800) used along with ATM shows a lot of radial distortion as it is not specifically meant for mapping purpose. If Tsai's calibration method is enhanced to calibrate up to fifth order distortion coefficients then the results would improve considerably.

- For rectifying the mounting and attitude errors Ground Control Points (GCP) are required. This is usually done by identifying noticeable target points in the surveyed images. A ground crew then gets the coordinates of the target points. The spot elevation data passing through the level1 of the LaserMap processing system has an accuracy of within a meter along X, Y and Z directions. This gridded data can be used to identify some target points to be used as GCPs, building corners and road intersections are a few such examples.

REFERENCES

[1] Abdel-Aziz Y.I., Karara H.M., "Direct Linear Transformation into Object Space Coordinates in Close-Range Photogrammetry". *Proc. ASP Symposium on Close-Range Photogrammetry, Urbana, USA*, pp. 1-18, 1971.

[2] Berg M., Kreveld M., Overmars M., Schwarzkopf O., *Computational Geometry Algorithms and Applications,* Springer Publications, 2nd Edition, 1998.

[3] Brock J., Sallenger A., "Airborne Topographic LIDAR Mapping for Coastal Science and Resource Management.", *U.S. Geological Survey Open-File Report*, No. 01-46, 2001.

[4] Brock J., Sallenger A., Krabill W., Swift R., Manizade S., Meredith A., Jansen M., Eslinger D., "Aircraft Laser Altimetry for Coastal Process Studies". *Coastal Sediments* pp.2414-2428, 1999.

[5] Brock J., Wright W., Sallenger A., Krabill W., Swift R., "Basis and Methods of NASA Airborne Topographic Mapper LIDAR Surveys for Coastal Studies". *Manuscript Submitted to the Journal of Coastal Research*, May 2001.

[6] Brown, D.C., "Close-Range Camera Calibration." *Photogrammetric Engineering*, vol 37, No. 8, August 1971, pp. 855-866.

[7] Brown, D.C., "Decentering Distortion of Lenses." *Photogrammetric Engineering*, Vol. 32, No. 3, pp. 444-462, 1966.

[8] Chen, T., "High Precision Georeference for Airborne Three-Line Scanner (TLS) Imagery." *Proceedings of* 3^{rd} *International Image Sensing Seminar on New Development in Digital Photogrammetry*, Gifu, Japan, pp. 71-82, 2001.

[9] Cormen T., Leiserson C., Rivest R., *Introduction to Algorithms,* McGraw Hill Publishing, 1997.

[10] Cramer, M., Stallmann, D. and Haala, N., "Direct Georeferencing Using GPS/Inert ial Exterior Orientations for Photogrammertric Applications." *International Archives of Photogrammetry and Remote Sensing*, Vol. 33 Part B3, pp. 198-205, 2000.

[11] Dorota A. Grejner-Brzezinska, "Direct Exterior Orientation of Airborne Imagery with GPS/INS System: Performance Analysis". *Journal of the Institute of Navigation*, Vol. 46, No. 4, pp. 261-270, 1999.

[12] Ebner, H., "Self-calibrating block adjustment." *Congress of the International Society for Photogrammetry, Invited Paper of Commission III*, Helsinki, Finland, 1976.

[13] Godard Space Flight Center, *http://www.gsfc.nasa.gov*, October 2002.

[14] Heikkilä J., "Geometric Camera Calibration using Circular Control Points". *IEEE Transactions on Pattern Analysis and Machine Intelligence.*, Vol. 22, No. 10, Oct. 2000.

[15] Heikkilä J., Silven O., "A Four-Step Camera Calibration Procedure with Implicit Image Correction". *Conference on Computer Vision and Pattern Recognition (CVPR'97)*, p. 1106-1112, June 17-19, 1997

[16] Hill J., Graham L., Henry R., Cotter D., Ding A., Young D., "Wide-Area Topographic Mapping and Applications Using Airborne Light Detection and Ranging (LIDAR) Technology". *Journal of the American Society for Photogrammetry and Remote Sensing*, Vol. 66, No. 8, 2000.

[17] Hohle, J., "Automatic Georectification of Aerial Images by Means of Existing Orthoimages and Height Data." *Proceedings of the ISPRS Commission II Symposium "Data Integration Systems and Techniques."*, Cambridge, England, pp. 121-126, 1998.

[18] Jain A. K., *Fundamentals of Digital Image Processing*, PHI Publishing, p. 384, 1997.

[19] Krabill W., "Airborne Topographic Mapper: An Introduction". *http://aol.wff.nasa.gov/aoltm.html*, August 2001.

[20] Krabill, Collins, Link, Swift, Butler, "Airborne Laser Topographic Mapping Results". *Photogrammetric Engineering & Remote Sensing*, Vol. 50, n6, pp. 685-694, 1984.

[21] Krabill W., Martin C., "Aircraft Positioning Using Global Positioning System Carrier Phase Data". *Navigation*, Vol. 34, pp.1-21, 1987.

[22] Krabill W., Wright W., Swift R., Frederick E., Manizade S., Yungel J., Martin C., Sonntag J., Duffy M., Hulslander W., Brock J., "Airborne Laser Mapping of Assateague National Seashore Beach". *Photogrammetric Engineering & Remote Sensing*, Vol.66, No.1, January 2000.

[23] Kratky V., "Rigorous Photogrammetric Processing of SPOT Images at CCM Canada". *ISPRS Journal of Photogrammetry and Remote Sensing*, No. 44, pp. 53-71, 1989.

[24] Lee, C., Theiss, H.J., Bethel, J.S., Mikhail, E.M., "Rigorous Mathematical Modeling of Airborne Pushbroom Imaging Systems." *Photogrammetric Engineering and Remote Sensing*, Vol. 66, No. 4, pp. 385-392, 2000.

[25] Moore L., "Shoreline Mapping Techniques". *Journal of Coastal Research*, Vol. 16, No. 1, pp.111-124.

[26] Nayegandhi, A., "Lasermap: Software System for Processing Topographic LIDAR Imagery." *Master's Thesis.*, 2001.

[27] NGDC/NOAA Shoreline/Coastline data, *http://www.ngdc.noaa.gov/mgg/shorelines/shorelines.html*, October 2002.

[28] Poli, D., "Direct Georeferencing of Multi-line Images with a General Sensor Model." *ISPRS Workshop "High resolution Mapping from Space 2001", Hanover, Germany.*, Sep. 2001.

[29] Poli, D., "Indirect Georeferencing of Airborne Multi-Line Array Sensors: A Simulated Case Study." *Proceedings of ISPRS Commission III Symposium "Photogrammetric Computer Vision '02", Graz, Austria.*, Vol. 34, No. B3/A, pp. 246-251.

[30] Schenk, T., "A Broader View on the Orientation Problem." *Presentation given at ISPRS workshop on "Direct versus Indirect methods of sensor Orientation."*, Barcelona, Spain, 1999.

[31] Schwarz, K.P., "Aircraft Position and Attitude Determinations by GPS and INS." *International Archives of Photogrammetry and remote Sensing.*, Vol. 31, Part B2, Vienna, pp. 67-73, 1996.

[32] Skaloud, J., Cramer, M., Schwarz, K.P., 1996, "Exterior Orientation by Direct Measurements of Camera Position and Attitude". *XVII. ISPRS Congress, Vienna, Austria, July 9-19, Int. Archives of Photogrammetry and Remote Sensing* vol 31, part B3, pp. 125-130, 1996.

[33] Sonka M., Hlavac V., Boyle R., *Image Processing, Analysis, and Machine Vision*, PWS Publishing, California, CA, p. 77, 233, 1999.

[34] Tsai R. Y., "An Efficient and Accurate Camera Calibration Technique for 3D machine Vision". *Proceedings of International Conference on Computer Vision and Pattern Recognition, Miami Beach, USA*, pp. 364-374, 1986.

[35] U.S. Geological Survey, "The Universal Transverse Mercator (UTM) Grid".*Fact Sheet 157-99. Available online at http://mac.usgs.gov/mac/isb/pubs/factsheets/fs15799.html,* September, 2001.

[36] Wallops Flight Facility, *http://www.wff.nasa.gov*, Jan 2002.

[37] Wehr A., Lohr U., "Airborne Laser Scanning - an Introduction and Overview". *ISPRS Journal of Photogrammetry & Remote Sensing*, Vol.54, pp.68-82, 1999.

[38] Zhang Z., "A Flexible New Technique for Camera Calibration". *IEEE Transactions on Pattern Analysis and Machine Intelligence.*, Vol. 22, No. 11, Nov. 2000.